DOCUMENT RESUME

ED 074 086 TM 002 448

AUTHOR	Harris, Chester W.
TITLE	Methodological Problems Encountered in the Project. Symposium: The Structure of Concept Attainment Abilities Project. Final Report and Critique.
PUB DATE	Feb 73
NOTE	8p.; Paper presented at the annual meeting of the American Educational Research Association, New Orleans, Louisiana, February 1973
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Achievement Tests; *Cognitive Tests; *Concept Formation; Factor Analysis; Item Analysis; *Research Problems; Scoring; *Test Construction
IDENTIFIERS	*Concept Attainment Abilities Project

ABSTRACT

Four problems of data summary and analysis in the Concept Attainment Abilities Project are discussed. The problems are: (1) scoring and item analysis for items that exist in a completely crossed design; (2) determining the factorial structure of such item sets; (3) reducing the battery of 56 cognitive abilities tests administered in 1970 to approximately 30 tests; and (4) identifying relationships between the battery of concept attainment measures and the battery of cognitive abilities tests. (DB)

FILMED FROM BEST AVAILABLE COPY

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 074086

TM 002 448

Symposium: The Structure of Concept Attainment Abilities Project
Final Report and Critique

METHODOLOGICAL PROBLEMS ENCOUNTERED IN THE PROJECT

Chester W. Harris
University of California, Santa Barbara

A paper presented at the annual meeting of the
American Educational Research Association
February 26, 1973
New Orleans

I shall comment briefly on four problems of data summary and analysis.

The construction and analysis of the concept attainment tests presented two problems. For each subject matter, a set of 360 items exists in a completely crossed design; the items form a matrix of, say, twelve rows representing the twelve tasks and thirty columns representing the thirty concepts. For a given individual who responds to all 360 items, this matrix can be used to represent his performance by filling in each cell with a "0" for an incorrect and a "1" for a correct response. The problem of how to use such student response data to improve items or to select the better items arises; this problem is not independent of the problem of how to score such a set of items.

One method of dealing with the 360 possible responses would be to assume that all responses indicate the same achievement and simply sample, at random, twenty or thirty of these responses and add the scores to yield a single score describing concept attainment for that student in that subject matter. This would permit forming indefinitely many randomly parallel tests that could be administered to different individuals and/or to the same individual at different times. However, such a procedure rests on the assumption that the row headings and the column headings do not designate important functional differences. This is an assumption we wished to test, and so we used redundant sets of scores. For concepts, thirty scores were developed by summing the columns of the task-concept matrix, and, for tasks, twelve scores were developed by summing the rows. We recognize that these two sets of scores are not completely independent. Also, only in the absence of concept-task interactions do the two sets of scores reflect the total variance in the matrix. This redundant scoring was employed in conventional item analysis to yield, for each item, two criterion scores. We found no simple and satisfactory way to put the two sets of item

parameter estimates together, and so we examined them separately. We feel that problems of scoring and of item analysis for items that exist in a completely crossed design are worthy of further investigation and study. Certainly we did not resolve apparent issues or solve these problems. It may be that test developers will discover more and more occasions on which the complete crossing of content and task is appropriate; if so, these problems of scoring and item analysis will deman solution.

A second problem associated with the measurement of concept attainment was that of determining the factorial structure of such item sets. It was clear that for any one subject matter, a three-mode analysis was needed to describe the task dimensions, the concept dimensions, and the concept-task interactions. However, we found the current programs limited in two respects. For one, most work has been dome on a component-type analysis, as contrasted with a factor-type analysis. We would like to see more capability for the latter. Second, there are, currently, rather sharp limits on the product of the number of instances in each of the two modes other than individuals. We reduced this product in two ways, by summing across the ten concepts within a sub-area and by summing across tasks. Both of these analyses tended to indicate no substantial concept-task interactions for any of the four subject matters. We recognize that the methods we were able to use here gave only partial answers to the question of content-task interaction, since in both analyses headings on one of the dimensions were aggregated.

We also performed conventional factor analyses for the two sets of scores for each of the four subject matter fields for both the 1970 and the 1971 data. Our intent was to apply the strategy for factor interpretation that has been described elsewhere (Harris & Harris, 1971), and so we secured

3 initial and 6 derived solutions for each matrix. The derived solutions tended to differ markedly in the number of common factors, with Alpha yielding the smallest number (usually one common factor) and Harris R-S^2 the largest number with one exception. This wide discrepancy made the application of our factor interpretation strategy difficult. Apparently our strategy, which involves identifying comparable common factors, is most appropriate when the derived solutions tend to have approximately the same number of major common factors. This was true for the analyses of the combined concept attainment measures and of the cognitive abilities tests, and for these analyses the strategy worked well. But within a subject matter field, the disparity forced us to consider whether or not the solutions with several factors were overdifferentiating. Such a condition is suggested when the oblique solutions have very high factor intercorrelations. This was found to be true, and we concluded that a single common factor characterizing the achievement scores of a subject matter was the most defensible interpretation. This interpretation was found to be consistent with the later results in which achievement measures for the four subject matter fields were analyzed together.

A third problem was that of reducing the battery of 56 cognitive abilities tests administered in 1970 to approximately 30 tests, without losing major factors that had been identified. Generally, one looks at the factor matrix, tries to find two variables that are of complexity one for each factor and chooses those to retain. Often--especially for orthogonal derived solutions--it is not possible to identify two variables that are essentially of complexity one for a given factor, and it then is necessary to "compromise." Our decisions were made primarily on the basis of the comparable common factors and if a variable were a possible candidate for inclusion, we relied heavily on the regression coefficients (pattern matrix) of the oblique solutions in deciding whether or not to choose it. It is clear that these

regression coefficients "go the other way" so to speak; that is, these pattern values are regression coefficients for estimating variables from factors, and not vice versa.

If we regard this problem of reducing a battery in such a way as to retain each of the major factors as a problem similar to that of estimating factor scores, then one might choose from several methods of developing the weights on the variables to use in estimating factors or factor scores. See Harris (1967) and Tucker (1971) for relevant analyses. Presumably those variables with substantial weights for estimating a given factor would be the ones from which to select--providing of course that the variable has only trivial weights for estimating any other factor. A first problem is that of selecting the system of developing the weights. We were influenced by Tucker's analysis that shows that if the factors are to be related to variables outside the set of variables on which the factor analysis was done, then the "ideal variable" or "least squares" weighting rpocedure is a good one. If we accept this, then oblique solutions which have been derived using the Harris-Kaiser independent cluster algorithm have pattern matrices which are proportional to the weights for estimating factors. Harris (1971) has commented on this elsewhere in connection with work done by Brogden (1969). The conclusion is to use this independent cluster model solution pattern matrix as the relevant matrix to examine and to choose variables which are of complexity one in that matrix. Our problem was not solved that easily, however, since we had found the independent cluster algorithm gave a certain amount of bipolarity in the pattern matrix, and we had substituted the Harris-Kaiser A'A proportional to L solution. In using the pattern matrix for this type of solution as a guide to the choice of variables to retain in the second study, we in effect assumed that the inverse of A'A was well

approximated by a diagonal matrix. It should be emphasized that this problem of how to reduce a battery and yet retain the factors that have been identified is not a trivial one.

The fourth problem I shall mention is that of identifying relationships between the battery of concept attainment measures and the battery of cognitive abilities tests. Different attitudes may be adopted in planning a study of the relations between two batteries; we adopted an interbattery approach. In contrast, some workers do not make a distinction between the two batteries and others approach the problem either by way of canonical correlation and canonical variates when a component-type solution is of interest, or by way of projecting the variables of the "dependent" battery into the common factor space of the "independent" battery when a factor-type solution is wanted.

We used canonical correlation theory for two purposes: to derive the squared multiple correlations for estimating each variable in one battery from the set of variables in the other battery, and to determine the number of significant canonical correlations to give us a guide to the number of interbattery factors to be extracted. At this concluding stage of the project, we wished to be certain that our generalizations about the dimensionality of the relations between concept attainment and cognitive abilities would reflect concern about the number of degrees of freedom available to us.

Our experience indicated that the Tucker interbattery procedure (Tucker, 1958) could readily be adapted to our type of study. The adaptations consisted of using a different test for the number of factors, securing derived orthogonal solutions for the two sets of interbattery factors that employed the same orthonormal transformation, and applying the Harris-Kaiser procedure for securing oblique solutions to the interbattery factors. The three adaptations

seemed to work well and may be considered by others. In general we found that the interbattery procedure is especially useful in that such factors are easy to extract and are easy to put into readily interpretable form. Our recommendation would be that if the derived solutions are to be secured by applying the same orthonormal solution (or the same oblique transformation), the transformation matrix should be developed from both sets of factors simultaneously. This seems likely to give more interpretable results, especially when there is a substantial number of variables in the smaller battery. We are speaking here, of course, of studies such as ours in which the two batteries are conceptually different, rahter than being interchangable, and the interest is in linking the two.

Our experience also indicated that the interbattery procedure gave more useful results than a conventional single solution involving both batteries simultaneously. Both the relative number of variables and the general level of correlations in the two batteries may influence the single solution; the interbattery solution is not influenced by the intrabattery correlation level, and may be relatively independent of the split of the variables into the two batteries.

We also developed the squared multiple correlations associated with estimating variables of one battery from the interbattery factor scores. (These factor scores are implicit in the theory, but are not uniquely calculable. The estimation procedure, however, does not depend on actually having these factor scores available.) We believe that these squared multiple correlations should have relatively little shrinkage in a new study from a similar population of students since we used the significance test to establish a (possibly) conservative number of dimensions for estimating the cross-correlations. If so, these squared multiple correlations should be good guides to future expectations.

REFERENCES

Brogden, H. E. Pattern, structure, and the interpretation of factors. Psychological Bulletin, 1969, 72, 375-378.

Harris, C. W. On factors and factor scores. Psychometrika, 1967, 32, 363-379.

Harris, C. W. On Brogden's interpretation of factors. Psychological Bulletin, 1971, 75, 360-361.

Harris, M. L., and Harris, C. W. A factor analytic interpretation strategy. Educational and Psychological Measurement, 1971, 31, 589-606.

Tucker, L. R. An inter-battery method of factor analysis. Psychometrika, 1958, 23, 111-136.

Tucker, L. R. Relations of factor score estimates to their use, Psychometrika, 1971, 36, 427-436.